BRL-TR-3303

AD-A243 507

TECHNICAL REPORT BRL-TR-3303

BRL

DATA FUSION FOR LEAST SQUARES

ANDREW ANDERSON THOMPSON III

DECEMBER 1991

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

U.S. ARMY LABORATORY COMMAND

BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND

NOTICES

UNCLASSIFIED

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1991	Final, January 1990—August 1991

4. TITLE AND SUBTITLE
Data Fusion for Least Squares

5. FUNDING NUMBERS
PR: 1L162618AH80

6. AUTHOR(S)
Andrew Anderson Thompson III

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Ballistic Research Laboratory
ATTN: SLCBR-DD-T
Aberdeen Proving Ground, MD 21005-5066

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
BRL-TR-3303

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The wide array of uses for least-squares estimation testifies to its effectiveness. The key to structuring a problem for a least-squares solution is finding a Markov representation of the problem. This representation defines a recursive approach to estimation. When multiple estimates are available at each time step, the processing time can be decreased by using data-fusion networks to reduce the information to a single estimate. Hierarchical networks, using both parallel and serial combination of data, can be devised. The cases presented herein can be used to preprocess the data for any recursive least-squares method.

14. SUBJECT TERMS
Data Fusion, Least Squares Method, Recursive Least Squares
Recursive Functions

15. NUMBER OF PAGES
33

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

UNCLASSIFIED

INTENTIONALLY LEFT BLANK.

TABLE OF CONTENTS

INTENTIONALLY LEFT BLANK.

ACKNOWLEDGMENT

The author wishes to thank David Webb, Joseph Olah, and Richard Sandmeyer for providing comments that improved this report.

INTENTIONALLY LEFT BLANK.

1. INTRODUCTION

As a decision maker, a person or a machine bases decisions on information received from the environment. Often, information pertinent to a specific decision comes from many sources. The purpose of this report is to develop the formulas to be used for linearly combining estimates. It is conceived that the estimates are being generated by some form of measurement where the uncertainty can be described as a Gaussian function. Examples of this type of measurement can include sensor systems, experimentation, or value judgements. To use these methods, the uncertainty associated with each source is assumed to be known. This report starts with the simplest situation and then looks at increasingly complicated fusion problems.

As an example, consider an active protection system for a tank. The tank may have several sensor systems that estimate properties of an incoming projectile. Interferometers, range sensors, and velocity sensors can be combined to give several independent estimates of a projectile's position at a specific time. Each of these estimates should be combined into a single improved estimate of position. This improved position estimate is then used to estimate the projectile's trajectory.

Least-squares estimation selects the values of the parameters in a mathematical model that minimize the squared differences between the mathematical model and a set of observations. Data fusion methods are based on a priori knowledge of a source's measurement error and result in a weighted average of the observations. Three general data processing approaches to estimation are "en bloc," recursive, and iterative estimation. The iterative approach calls for multiple passes through the data and will not be discussed in this paper. In the "en bloc" method, all the data is processed at once to calculate the estimate. When the weights are calculated "en bloc," each weight indicates the relative value of the source. The recursive method, where the data is processed one observation at a time, develops the combination rule used in Kalman filtering. Although there is a statistical problem with the recursive method in the initialization of the process, the recursive formulation of the least-squares solution seems to be preferred in many fields, including electronics, economics, and biology. The recursive least-squares method has several pragmatic advantages over the "en bloc" method in real time cases. These include:

1. An estimate and its error distribution are always available.

2. A decision can be made based on the current error distribution.

3. It can provide more insight to the actual problem.

4. It can be modified into different approximation techniques when the underlying least-squares assumption are compromised.

In recursive estimation, the estimate is updated each time an observation becomes available. The weight associated with the observation indicates the value of the observation in relation to the value of the current estimate. The change in the estimate as a result of the update is called the gradient.

A fundamental process in data fusion is to find a representation of the target, or unknown system, so that updates based on new information depend only on the current estimate and the new data. This characteristic is referred to as the Markov property. In some situations, a list of resources, a deployment pattern, some doctrinal procedures, and a location may constitute a Markov representation. When the new information becomes available, the estimate is changed by a gradient that reflects the uncertainty associated with both the estimate and new observation. The ideas discussed herein can be used to determine the proper gradient when the uncertainty is known.

Consider the intelligence problem of a command center. Information must be combined and processed to identify enemy locations, type of units, identity of units, and the intentions or orders of the unit. Sources of information are reports from imagery intelligence, signal intelligence, and human intelligence. The value of each report depends on its source and its timeliness. As new information comes in, the current target estimate must be updated. Each report can be thought of as producing a gradient. When mathematical models of the target are available, it is possible to develop automated methods resulting in the best update. Even when automated methods are not possible, it is usually desirable to find a Markov representation and use this type of reasoning.

The report progresses on a case by case basis. The first seven cases deal with "en bloc" procedures while the last two methods are recursive. Case One considers the problem of estimating a single parameter from two noisy estimates. The basic method for solving the problem is demonstrated, and the formula for combining information is given. Case Two discusses the modifications to the basic formula if the uncertainty is a function of some variable (time or range for example). Case Three and Case Four introduce correlation between observations to the first two cases. Cases Five and Six extend the ideas to three observations and present the solution in the form of the general solution. Case Seven presents the general "en bloc" solution as a summary of the previous situations. Case Eight presents the recursive method for solving this type of problem. Case Nine introduces the recursive method for combining vector estimates.

2. CASE ONE

The method described here is for combining two different uncorrelated pieces of information from different sources. The quantity to be estimated is X; the goal is to find the form of the estimator $\hat{X}$.

2.1 Problem. Find the best way to linearly combine two observations, Z_1 and Z_2, if

$$Z_1 = X + V_1, \qquad V_1 \sim N(0, \sigma^2_1),$$

$$Z_2 = X + V_2, \qquad V_2 \sim N(0, \sigma^2_2), \text{ and}$$

$$E(V_1\ V_2) = 0 .$$

The estimator will have the form $\hat{X} = k_1 Z_1 + k_2 Z_2$.

2.2 Solution. If the estimator is to be unbiased, then

$$E(\hat{X} - X) = 0$$

or

$$E(k_1 Z_1 + k_2 Z_2 - X) = 0$$

$$E(k_1 X + k_1 V_1 + k_2 X + k_2 V_2 - X) = 0$$

$$k_1 X + k_2 X - X + k_1 E(V_1) + k_2 E(V_2) = 0$$

$$k_1 + k_2 - 1 = 0$$

$$k_1 = 1 - k_2 .$$

Since $k_1 + k_2 = 1$, we can simplify the notation by letting $k_1 = k$ and $k_2 = 1 - k$. After doing this, the form of the estimator is

$$\hat{X} = kZ_1 + (1 - k)Z_2 .$$

The variance of this estimator is found as follows:

$$\begin{aligned} E(\hat{X} - X)^2 &= E(kZ_1 + (1 - k)Z_2 - X)^2 \\ &= E(kX + kV_1 + (1 - k)X + (1 - k)V_2 - X)^2 \\ &= E\left[(k + (1 - k) - 1)X + kV_1 + (1 - k)V_2\right]^2 \\ &= E(kV_1 + (1 - k)V_2)^2 \\ &= E(k^2 V_1^2 + (1 - k)^2 V_2^2 + 2k(1 - k)V_1 V_2) \\ &= k^2 \sigma^2_1 + (1 - k)^2 \sigma^2_2 + 0 . \end{aligned}$$

To find the minimum variance estimator, take the derivative of the variance with respect to k, set this expression equal to zero and solve for k.

$$\frac{\partial}{\partial k} E(\hat{X} - X)^2 = 2k\sigma^2_1 - 2(1 - k)\sigma^2_2 = 0$$

$$k\sigma^2_1 + k\sigma^2_2 = \sigma^2_2$$

$$k = \frac{\sigma^2_2}{\sigma^2_1 + \sigma^2_2} .$$

Finding the second derivative verifies that this value of k is a minimum.

The form of the estimator is

$$\hat{X} = \frac{\sigma^2_2}{\sigma^2_1 + \sigma^2_2} Z_1 + \frac{\sigma^2_1}{\sigma^2_1 + \sigma^2_2} Z_1 \, .$$

The variance of the estimator can then be found.

$$\mathrm{Var}(\hat{X}) = \left[\frac{\sigma^2_2}{\sigma^2_1 + \sigma^2_2}\right]^2 \mathrm{Var}\ (Z_1) + \left[\frac{\sigma^2_1}{\sigma^2_1 + \sigma^2_2}\right]^2 \mathrm{Var}(Z_2) + 2\frac{\sigma^2_1\sigma^2_2}{(\sigma^2_1 + \sigma^2_2)^2} E(Z_1 Z_2)$$

$$= \frac{\sigma^4_2\sigma^2_1}{(\sigma^2_1 + \sigma^2_2)^2} + \frac{\sigma^4_1\sigma^2_2}{(\sigma^2_1 + \sigma^2_2)^2}$$

$$= \frac{\sigma^2_1\sigma^2_2(\sigma^2_2 + \sigma^2_1)}{(\sigma^2_1 + \sigma^2_2)^2}$$

$$= \frac{\sigma^2_1\sigma^2_2}{\sigma^2_1 + \sigma^2_2} \, .$$

To summarize, the estimate is

$$\hat{X} = \frac{\sigma^2_2}{\sigma^2_1 + \sigma^2_2} Z_1 + \frac{\sigma^2_1}{\sigma^2_1 + \sigma^2_2} Z_2$$

with a variance of

$$\frac{\sigma^2_1\sigma^2_2}{\sigma^2_1 + \sigma^2_2} \, .$$

3. CASE TWO

In this section, the results of the previous case are extended to consider the situation where the uncertainty has a functional form. For example, the uncertainty of a measurement may be a function of the magnitude of the measurement. This is the case for radars measuring range and for value judgements. Also, the uncertainty associated with some information may increase

over time. This happens in Kalman filtering when the current estimate is propagated forward in time as a prediction based on the state model.

If the variance of the estimator is a function of a variable, then the results of Case One can be rewritten as follows.

Assume

$$Z_1 = X + V_1, \qquad V_1 \sim N(0, f(R)),$$
$$Z_2 = X + V_2, \qquad V_2 \sim N(0, f(t)), \text{ and}$$
$$E(V_1 V_2) = 0 .$$

Then,

$$\hat{X} = \frac{f(t)}{f(t) + f(R)} Z_1 + \frac{f(R)}{f(t) + f(R)} Z_2$$

and

$$\mathrm{Var}\,(\hat{X}) = \frac{f(t)\, f(R)}{f(t) + f(R)} = \frac{1}{\frac{1}{f(t)} + \frac{1}{f(R)}}$$

Consider the following applications:

- **Example 1.**

Find the properties of the best estimator for two sensors if one has a constant standard deviation of 1 m and the other has a standard deviation of .05R where R is the range in meters; i.e.,

$$V_1 \sim N(0, 1^2)$$
$$V_2 \sim N(0, (.05R)^2) .$$

By the above results,

$$\hat{X} = \frac{(.05R)^2}{1 + (.05R)^2} Z_1 + \frac{1}{1 + (.05R^2)} Z_2, \text{ and}$$

$$\mathrm{Var}\,(\hat{X}) = \frac{(.05R)^2}{1 + (.05R)^2} = \frac{1}{1 + \frac{1}{(.05R)^2}} .$$

It can be seen that when R is large the variance approaches one and when R is small then it goes to $(.05R)^2$ (and then to zero as R goes to zero).

- **Example 2.**

Find the properties of an estimator that combines information whose variance increases exponentially with time; i.e.,

$$V_1 \sim N(0, e^{t_1}\sigma_1^2)$$

and

$$V_2 \sim N(0, e^{t_2}\sigma_2^2).$$

If we assume the estimates were made at t_1 and t_2 time units in the past, then

$$\hat{X} = \frac{e^{t_2}\sigma_2^2}{e^{t_1}\sigma_1^2 + e^{t_2}\sigma_2^2} Z_1 + \frac{e^{t_1}\sigma_1^2}{e^{t_1}\sigma_1^2 + e^{t_2}\sigma_2^2} Z_2 ,$$

and

$$\mathrm{Var}(\hat{X}) = \frac{e^{t_2}\sigma_2^2 e^{t_1}\sigma_1^2}{e^{t_1}\sigma_1^2 + e^{t_2}\sigma_2^2}$$

$$= \frac{e^{t_2+t_1}\sigma_1^2\sigma_2^2}{e^{t_1}\sigma_1^2 + e^{t_2}\sigma_2^2} .$$

In many situations, the results of Case One can be extended to find a quantitative data-fusion technique. The variations will depend on specific knowledge of the situation.

4. CASE THREE

Case Three extends Case One to consider the effects of correlated errors. Sometimes the information sources are not independent and the errors associated with each contain some common components. If we know the amount of association, the form of the estimator can be derived.

For two variables with correlated noise, the assumptions are:

$$Z_1 = X + V_1, \qquad V_1 \sim N(0, \sigma^2_1),$$

$$Z_2 = X + V_2, \qquad V_2 \sim N(0, \sigma^2_2), \text{ and}$$

$$E(V_1 V_2) = \rho \sigma_1 \sigma_2 .$$

As in Case One

$$\hat{X} = kZ_1 + (1 - k)Z_2 .$$

Under the assumptions, the variance of the estimator is

$$E(\hat{X} - X)^2 = E\,(k^2 V^2_1 + (1 - k)^2 V^2_2 + 2k\,(1 - k)\,V_1 V_2)$$

$$= k^2 \sigma^2_1 + (1 - k)^2 \sigma^2_2 + 2k\,(1 - k)\,\rho \sigma_1 \sigma_2 .$$

The minimum variance estimator is found as follows.

$$\frac{\partial}{\partial k} E(\hat{X} - X)^2 = 2k\sigma^2_1 - 2(1 - k)\sigma^2_2 + 2\rho\sigma_1\sigma_2 - 4k\rho\sigma_1\sigma_2$$

$$= (2\sigma^2_1 - 4\rho\sigma_1\sigma_2 + 2\sigma^2_2)k - 2\sigma^2_2 + 2\rho\sigma_1\sigma_2 .$$

After setting the partial equal to zero and solving for k, we get

$$k = \frac{\sigma^2_2 - \rho\sigma_1\sigma_2}{\sigma^2_1 + \sigma^2_2 - 2\rho\sigma_1\sigma_2} .$$

Putting this value into the above variance expression, we have the following steps to arrive at a simplified variance expression.

$$\mathrm{Var}\,(\hat{X}) = k^2\sigma_1^2 + (1 - k)^2\sigma_2^2 + 2k(1 - k)\,\rho\,\sigma_1\sigma_2$$

$$= \frac{(\sigma_2^2 - \rho\,\sigma_1\sigma_2)^2\sigma_1^2 + (\sigma_1^2 - \rho\,\sigma_1\sigma_2)^2\sigma_2^2 + 2(\sigma_2^2 - \rho\,\sigma_1\sigma_2)(\sigma_1^2 - \rho\,\sigma_1\sigma_2)\,\rho\,\sigma_1\sigma_2}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2(\sigma_2^2 - 2\rho\,\sigma_1\sigma_2 + \rho^2\sigma_1^2) + \sigma_1^2\sigma_2^2(\sigma_1^2 - 2\rho\,\sigma_1\sigma_2 + \rho^2\rho_2^2)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$+ \frac{2\rho\,\sigma_1\sigma_2(\sigma_1^2\sigma_2^2 - \rho\,\sigma_1^3\sigma_2 - \rho\,\sigma_1\sigma_2^3 + \rho^2\sigma_1^2\sigma_2^2)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2\left[\sigma_1^2 + \sigma_2^2 + \rho^2\sigma_1^2 + \rho^2\sigma_2^2 - 4\rho\,\sigma_1\sigma_2 + 2\rho\,\sigma_1\sigma_2 - 2\rho^2\sigma_1^2 - 2\rho\,\sigma_2^2 + 2\rho^3\sigma_1\sigma_2\right]}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2\left[\sigma_1^2 + \sigma_2^2 - \rho^2\sigma_1 - \rho^2\sigma_2 - 2\rho\,\sigma_1\sigma_2 + 2\rho^3\sigma_1\sigma_2\right]}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2(1 - \rho^2)(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2)^2}$$

$$= \frac{\sigma_1^2\sigma_2^2(1 - \rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2}.$$

In summary, the estimator has the form

$$\hat{X} = \frac{\sigma_2^2 - \rho\,\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2}Z_1 + \frac{\sigma_1^2 - \rho\,\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2}Z_2$$

with variance

$$\mathrm{Var}\,(\hat{X}) = \frac{\sigma_1^2\sigma_2^2(1 - \rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\,\sigma_1\sigma_2}.$$

5. CASE FOUR

This case presents the most general form of an estimator for combining two pieces of information. Case Three is extended to consider the case when the uncertainty associated with correlated information has a functional form. The variance is expressed as a function of a variable such as time, range, or degrees from boresight. With correlated observations, the form is similar to Case Two.

Assume

$$Z_1 = X + V_1, \qquad V_1 \sim N(0,f(r)),$$

$$Z_2 = X + V_2, \qquad V_2 \sim N(0,g(t)), \text{ and}$$

$$E(V_1 V_2) = h(r,t).$$

When combining two pieces of information, the most general form of the estimator is

$$\hat{X} = \frac{g(t) - h(r,t)}{f(r) + g(t) - 2h(r,t)} Z_1 + \frac{f(r) - h(r,t)}{f(r) + g(t) - 2h(r,t)} Z_2$$

with a variance of

$$\mathrm{Var}\,(\hat{X}) = \frac{f(r)g(t)\left[1 - \dfrac{h(r,t)^2}{g(t)f(r)}\right]}{f(r) + g(t) - 2h(r,t)}$$

6. CASE FIVE

In this case, three data points with uncorrelated measurement errors are combined as a single estimate. This extends Case One so that three pieces of information can be processed at one time. The assumptions for $i = \{1,2,3\}$ are

$$Z_i = X + V_i, \qquad V_i \sim N(0, \sigma^2{}_1),$$

$$E(V_i V_j) = 0 \text{ if } i \neq j.$$

Using the same reasoning as in Case One, it can be shown that $1 = k_1 + k_2 + k_3$. Note that $k_3 = 1 - k_1 - k_2$; thus, k_3 can be eliminated. The form of the estimator is $\hat{X} = k_1 Z_1 + k_2 Z_2 + (1 - k_1 - k_2) Z_3$. To find the values of k_1 and k_2 which minimize the variance, the partial derivatives of Var $(\hat{X})$ are found and set equal to zero. This set of equations is then solved for k_1 and k_2. The following equations realize these steps.

$$\text{Var}\,(\hat{X}) = k^2{}_1\sigma^2{}_1 + k^2{}_2\sigma^2{}_2 + (1 - k_1 - k_2)^2\sigma^2{}_3$$

$$\frac{\partial}{\partial k_1}(\text{Var}\,(\hat{X})) = 2k_1\sigma^2{}_1 - 2(1 - k_1 - k_2)\sigma^2{}_3$$

$$\frac{\partial}{\partial k_2}(\text{Var}(\hat{X})) = 2k_2\sigma^2{}_2 - 2(1 - k_1 - k_2)\sigma^2{}_3$$

Setting the partials equal to zero, we have the following two equations:

$$(\sigma^2{}_1 + \sigma^2{}_3)k_1 + \sigma^2{}_3 k_2 = \sigma^2{}_3, \text{ and}$$

$$\sigma^2{}_3 k_1 + (\sigma^2{}_2 + \sigma^2{}_3)k_2 = \sigma^2{}_3.$$

Using matrix notation, these equations can be written as

$$\begin{bmatrix} \sigma^2{}_1 + \sigma^2{}_3 & \sigma^2{}_3 \\ \sigma^2{}_3 & \sigma^2{}_2 + \sigma^2{}_3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \sigma_3 \\ \sigma_3 \end{bmatrix}.$$

The above matrix equation can be rewritten as

$$\begin{bmatrix} \dfrac{\sigma_1^2}{\sigma_3^2} + 1 & 1 \\ 1 & \dfrac{\sigma_2^2}{\sigma_3^2} + 1 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

From this, the determinant can be seen to be

$$\det = \frac{\sigma_1^2 \sigma_2^2}{\sigma_3^4} + \frac{\sigma_1^2}{\sigma_3^2} + \frac{\sigma_2^2}{\sigma_3^2}$$

$$= \frac{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}{\sigma_3^4}.$$

By Cramer's rule, we find

$$k_1 = \frac{\sigma_2^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}, \text{ and } k_2 = \frac{\sigma_1^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}.$$

The variance of the estimator is

$$\operatorname{Var}(\hat{X}) = k_1^2 \sigma_1^2 + k_2^2 \sigma_2^2 + k_3^2 \sigma_3^2$$

$$= \frac{\sigma_2^4 \sigma_3^4 \sigma_1^2 + \sigma_1^4 \sigma_3^4 \sigma_2^2 + \sigma_1^4 \sigma_2^4 \sigma_3^2}{(\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2)^2}$$

$$= \frac{\sigma_1^2 \sigma_2^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2}$$

$$= \frac{1}{\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2} + \dfrac{1}{\sigma_3^2}}.$$

The variance of the estimator has the same form as the resistance of a parallel circuit. From an electrical engineering perspective, the ideas could be expressed as resistance to the flow of

information. Indeed, a circuit of parallel transistors would be able to simulate this situation. The gate voltage would be set to create the proper resistance and the current through the circuit would be the variance of the estimate. For N independent measurements, this can be generalized to

$$k_j = \frac{\prod_{i \neq j}^{N} \sigma_i^2}{\sum_{i=1}^{N} \prod_{k \neq i}^{N} \sigma_k^2}$$

In general, one would have the estimate

$$\hat{X} = \sum_{i=1}^{N} k_i Z_i$$

with a variance of

$$\mathrm{Var}\,(\hat{X}) = \frac{1}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}$$

In summary, for three observations the form of the estimator is

$$\hat{X} = \frac{\sigma_2^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2} Z_1 + \frac{\sigma_1^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2} Z_2 + \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2} Z_3 \,.$$

The variance of $\hat{X}$ is

$$\mathrm{Var}\,(\hat{X}) = \frac{\sigma_1^2 \sigma_2^2 \sigma_3^2}{\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2} \,.$$

7. CASE SIX

This case extends Case Five by including the effects of correlated noise. The matrix representation used in this and the previous case is needed for the general case presented in the next section. The assumptions for i = {1,2,3} are represented as follows

$$Z_i = X + V_i, \qquad V_i \sim N(0, \sigma^2_i),$$

$$E(V_i V_j) = \rho_{ij}\sigma_{ij} \text{ if } i \neq j .$$

As in Case Five, $1 = k_1 + k_2 + k_3$. Due to the correlation between observations, the variance expression is more complicated. It is

$$\begin{aligned}
\text{Var}\ (\hat{X}) &= k^2_1\sigma^2_1 + k^2_2\sigma^2_2 + k^2_3\sigma^2_3 + 2\rho_{12}k_1k_2\sigma_1\sigma_2 + 2\rho_{13}k_1k_3\sigma_1\sigma_3 \\
&\quad + 2\rho_{23}k_2k_3\sigma_2\sigma_3 \\
&= k^2_1\sigma^2_1 + k^2_2\sigma^2_2 + (1 - k_1 - k_2)^2\sigma^2_3 + 2\rho_{12}k_1k_2\sigma_1\sigma_2 \\
&\quad + 2\rho_{13}k_1(1 - k_1 - k_2)\sigma_1\sigma_3 + 2\rho_{23}k_2(1 - k_1 - k_2)\sigma_2\sigma_3 .
\end{aligned}$$

To find the minimum variance, the partial derivatives are set equal to zero and then the resulting matrix equation is solved.

$$\begin{aligned}
\frac{\partial}{\partial k_1}\text{Var}\ (\hat{X}) &= 2\sigma^2_1 k_1 - 2(1 - k_1 - k_2)\sigma^2_3 + 2\rho_{12}\sigma_1\sigma_2 k_2 \\
&\quad + 2\rho_{13}\sigma_1\sigma_3(1 - 2k_1 - k_2) - 2\rho_{23}\sigma_2\sigma_3 k_2 \\
&= 2(\sigma^2_3 + \rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3 - \rho_{23}\sigma_2\sigma_3)k_2 \\
&\quad + 2(\sigma^2_1 + \sigma^2_3 - 2\rho_{13}\sigma_1\sigma_3)k_1 \\
&\quad - 2(\sigma^2_3 - \rho_{13}\sigma_1\sigma_3) \\
&= 0
\end{aligned}$$

Similarly,

$$\frac{\partial}{\partial k_2} \mathrm{Var}\,(\hat{X}) = 2(\sigma_3^2 + \rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3 - \rho_{23}\sigma_2\sigma_3)\,k_1$$

$$+ 2(\sigma_2^2 + \sigma_3^2 - 2\rho_{23}\sigma_2\sigma_3)\,k_2$$

$$- 2(\sigma_3^2 - \rho_{23}\sigma_2\sigma_3)$$

$$= 0$$

In matrix representation, this is

$$\begin{bmatrix} \sigma_1^2 + \sigma_3^2 - 2\rho_{13}\sigma_1\sigma_3 & \sigma_3^2 + \rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3 - \rho_{23}\sigma_2\sigma_3 \\ \sigma_3^2 + \rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3 - \rho_{23}\sigma_2\sigma_3 & \sigma_2^2 + \sigma_3^2 - 2\rho_{23}\sigma_2\sigma_3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} =$$

$$\begin{bmatrix} \sigma_3^2 - \rho_{13}\sigma_1\sigma_3 \\ \sigma_3^2 - \rho_{23}\sigma_2\sigma_3 \end{bmatrix}.$$

The above can be solved by using Cramer's rule, or some other matrix algebra technique.

8. CASE SEVEN

The general case for finding the optimal set of weights $\{k_i\}$ is developed in this section. This is the last "en bloc" method discussed. The results are derived by finding the partial of the variance expression with respect to one variable (k_l) then, after setting this expression equal to zero and solving it, the proper matrix equation is shown. The approach taken is to break the expression into a series of terms. Each term is examined in sequence then these results are combined into the partial of the variance with respect to the selected variable.

The variance of the estimator for the general case is

$$\mathrm{Var}\,(\hat{X}) = \sum_{i=1}^{n} \sigma_i^2 k_i^2 + \sum_{i=1}^{n-1} \sum_{j>i}^{n} 2\rho_{ij}\sigma_i\sigma_j k_i k_j\,. \qquad (1)$$

Note that

$$k_n = \left(1 - \sum_{m=1}^{n-1} k_m \right).$$

To get the partial of the variance with respect to the k_l, $0 < l < n\text{-}1$, the partial of the first term of Equation (1) is found and then the partial of the second term. For the first term, we have

$$\begin{aligned}
\frac{\partial}{\partial k_l} \sum_{i=1}^{n} \sigma_i^2 k_i^{\,2} &= \frac{\partial}{\partial k_l} \sum_{i=1}^{n-1} \sigma_i^2 k_i^{\,2} + \frac{\partial}{\partial k_l} \sigma_n^2 \, (1 - \sum_{m=1}^{n-1} k_m)^2 \\
&= 2\sigma_l^2 k_l - 2\sigma_n^2 \, (1 - \sum_{m=1}^{n-1} k_m) \\
&= 2\sigma_l^2 k_l - 2\sigma_n^2 + 2\sigma_n^2 k_l + 2 \sum_{i \neq l}^{n-1} \sigma_n^2 k_i \, .
\end{aligned} \tag{2}$$

The second term is more complicated than the first term.

$$\frac{\partial}{\partial k_l} \sum_{i=1}^{n=1} \sum_{j>i}^{n} 2\rho_{ij} \sigma_i \sigma_j k_i k_j =$$

$$\frac{\partial}{\partial k_l} \sum_{i=1}^{n-2} \sum_{j>i}^{n-1} 2\rho_{ij} \sigma_i \sigma_j k_i k_j + \frac{\partial}{\partial k_l} \sum_{i=1}^{n-1} 2\rho_{in} \sigma_n \sigma_i k_i (1 - \sum_{m}^{n-1} k_m)$$

Each of the above terms will be considered (and referred to as the third and fourth term). In the third term, note that $i \neq j$; so, if the partial term is nonzero, either $i = l$ or $j = l$. For example, let $n = 5$ and $l = 3$, then if $i = 1$ and $j = 3$, a nonzero partial term is obtained; likewise, when $i = 2$ and $j = 3$, and when $i = 3$ and $j = 4$, nonzero partial terms are obtained. The partial of the third term can be written as follows:

$$\frac{\partial}{\partial k_l} \sum_{i=j}^{n-2} \sum_{j>i}^{n-1} 2\rho_{ij} \sigma_i \sigma_j k_i k_j = \begin{cases} \displaystyle\sum_{j=l+1}^{n-1} 2\rho_{lj} \sigma_l \sigma_j k_j, & \text{if } i = l \\ \displaystyle\sum_{i=1}^{l-1} 2\rho_{il} \sigma_i \sigma_l k_i, & \text{if } j = l \end{cases} .$$

Combining this result into a single expression, we have

$$\sum_{i \neq l}^{n-1} 2\rho_{il}\sigma_l\sigma_i k_i \tag{3}$$

as the partial derivative.

The fourth term will now be evaluated; i.e.,

$$\frac{\partial}{\partial k_l} \sum_{i=1}^{n-1} 2\rho_{in}\sigma_n\sigma_i k_i (1 - \sum_{m=1}^{n-1} k_m).$$

For exactly one value, i will equal l; also, for every value of i, one k_m will equal k_l. Consider the case $i \neq l$; recall that exactly one k_m will be k_l.

$$\frac{\partial}{\partial k_l} \sum_{i \neq l}^{n-1} 2\rho_{in}\sigma_n\sigma_i k_i (1 - \sum_{m=1}^{n-1} k_m) = \sum_{i \neq l}^{n-1} - 2\rho_{in}\sigma_n\sigma_i k_i \tag{4}$$

Consider the case $i = l$.

$$\frac{\partial}{\partial k_l} 2\rho_{ln}\sigma_n\sigma_l k_l (1 - \sum_{m=1}^{n-1} k_m) = 2\rho_{ln}\sigma_n\sigma_l (1 - \sum_{m=1}^{n-1} k_m - k_l)$$

$$= 2\rho_{ln}\sigma_n\sigma_l - 2\rho_{ln}\sigma_n\sigma_l \sum_{i \neq l}^{n-1} k_i - 4\rho_{ln}\sigma_l\sigma_n k_l. \tag{5}$$

The next task is to collect the various expressions (2, 3, 4, 5) that make up the partial derivative and then to organize them in some meaningful format. There will be three groupings; terms containing k_l, terms containing no variables, and terms containing variables other than k_l.

Terms associated with k_l are

$$2\sigma_l^2 k_l + 2\sigma_n^2 k_l - 4\rho_{ln}\sigma_l\sigma_n k_l.$$

Terms that do not contain a variable are

$$-2\sigma_n^2 + 2\rho_{ln}\sigma_l\sigma_n \, .$$

Terms associated with a variable k_i, where $i \neq l$ are

$$2\sum_{i \neq l}^{n-1} \sigma_n^2 k_i + 2\sum_{i \neq l}^{n-1} \rho_{il}\sigma_l\sigma_i k_i - 2\sum_{i \neq l}^{n-1} \rho_{in}\sigma_n\sigma_i k_i - 2\sum_{i \neq l}^{n-1} \rho_{ln}\sigma_l\sigma_n k_i \, .$$

The above represents the partial derivative of one variable; by doing this for each variable, setting the result equal to zero, and dividing by two we can formulate a matrix equation for the general case.

The indices i, j will be used in place of l and i. The form we wish is

$$AK = C$$

where K is the variable vector of length n – 1, C a vector of constant terms of length n – 1, and A the coefficient matrix of size (n – 1) x (n – 1) of the vector K.

The diagonal terms of A, (a_{ii}), will be

$$a_{ii} = \sigma_i^2 + \sigma_n^2 - 2\rho_{in}\sigma_i\sigma_n \, ,$$

and the off diagonal terms of A, (a_{ij}), where $i \neq j$ will be

$$a_{ij} = \sigma_n^2 + \rho_{ij}\sigma_i\sigma_j - \rho_{jn}\sigma_j\sigma_n - \rho_{in}\sigma_i\sigma_n \, .$$

The elements of the vector C, (c_i), are

$$c_i = \sigma_n^2 - \rho_{in}\sigma_i\sigma_n \, .$$

The values of K can be found using matrix algebra ($K = A^{-1}C$). As with the two observation cases, functions can be substituted for the variance terms to produce classes of variable estimates.

The matrix A above can also be generated using the covariance matrix by the following method:

Let Σ be the covariance matrix with

$$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j .$$

If we generate the following matrices

$$D = \begin{bmatrix} 1 & 0 & . & . & . & -1 \\ 0 & 1 & . & . & . & -1 \\ 0 & 0 & . & . & . & -1 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{bmatrix}_{(n-1) \times r} \qquad \Sigma = \begin{bmatrix} \sigma^2_1 & \sigma_{12} & . & . & . & . \\ \sigma_{12} & \sigma^2_2 & . & . & . & . \\ . & . & \sigma^2_3 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{bmatrix}_{n \times n}$$

$$D' = \begin{bmatrix} 1 & 0 & 0 & . & . & . \\ 0 & 1 & 0 & . & . & . \\ 0 & 0 & 1 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ -1 & -1 & -1 & . & . & . \end{bmatrix}_{n \times (n-1)} ,$$

then $A = D \, \Sigma \, D'$.

9. CASE EIGHT

In contrast to "en bloc" procedures, recursive methods process the data one piece at a time and keep updating the estimate. As an example, assume unassociated observations Z_1 and Z_2 become available.

Assume one piece of data is available, our first estimate will be $\hat{X}_1$. The process is initialized so that

$$\hat{X}_1 = Z_1, \quad \mathrm{Var}\,(\hat{X}_1) = \sigma^2_1 .$$

If observation Z_2 becomes available, then

$$\hat{X}_2 = \frac{\sigma^2_2}{\sigma^2_1 + \sigma^2_2}\,\hat{X}_1 + \frac{\sigma^2_1}{\sigma^2_1 + \sigma^2_2}\,Z_2,$$

with $\mathrm{Var}\,(\hat{X}_2) = \dfrac{\sigma^2_1\,\sigma^2_2}{\sigma^2_1 + \sigma^2_2}$.

Assume Z_3 becomes available, then

$$\hat{X}_3 = \frac{\sigma^2_3}{\mathrm{Var}\,(\hat{X}_2) + \sigma^2_3}\,\hat{X}_2 + \frac{\mathrm{Var}\,(\hat{X}_2)}{\mathrm{Var}\,(\hat{X}_2) + \sigma^2_3}\,Z_3,$$

with

$$\mathrm{Var}\,(\hat{X}_3) = \frac{\mathrm{Var}\,(\hat{X}_2)\,\sigma^2_3}{\mathrm{Var}\,(\hat{X}_2) + \sigma^2_3}.$$

It can be verified that this result agrees with Case Five.

Note that the number of steps used in the recursive solution is a linear function of the number of data points. (To update the estimate and its variance takes two additions, three multiplications, and two divisions.) The recursive method is usually preferred for real time applications and in control systems.

Recursive techniques express the estimate as a weighted average of the old estimate and the new observation. In many formulations, this is expressed as the old estimate plus the change as a result of the new information. The recursive estimator is:

$$\hat{X}_{i+1} = \frac{\sigma^2_{i+1}}{\mathrm{Var}\,(\hat{X}_i) + \sigma^2_{i+1}}\,\hat{X}_i + \frac{\mathrm{Var}\,(\hat{X}_i)}{\mathrm{Var}\,(\hat{X}_i) + \sigma^2_{i+1}}\,Z_{i+1}$$

which can be rewritten as:

$$\hat{X}_{i+1} = \hat{X}_i + \frac{\mathrm{Var}\,(\hat{X}_1)}{\mathrm{Var}\,(\hat{X}_i) + \sigma^2_{i+1}}\,(Z_{i+1} - \hat{X}_i)\,.$$

The second term represents the gradient due to the (i+1)st observation.

10. CASE NINE

To extend the idea of a recursive estimator to the vector situation is straightforward. Let $\hat{X}$ and Z_i represent vectors and replace the variance terms with the covariance matrices Σ_i that represent the uncertainties associated with each vector to be estimated. Note X X' is represented by X^2 and K is a matrix. The separate observations are assumed to be uncorrelated.

$$Z_1 = X + V_1\,, \qquad V_1 \sim N(0,\Sigma_1)\,,$$
$$Z_2 = X + V_2\,, \text{ and } V_2 \sim N(0,\Sigma_2)\,.$$

Using the same method as Case One, but with vectors

$$\begin{aligned} E(K_1 Z_1 + K_2 Z_2) &= E\,[K_1 X + K_1 V_1 + K_2 X + K_2 V_2] \\ &= K_1 X + K_2 X \\ &= (K_1 + K_2)\,X\,. \end{aligned}$$

If the estimator is to be unbiased, then

$$[K_1 + K_2]\, X = X .$$

Thus, $K_1 + K_2 = I$. The next example shows a method of solution when the two covariance structures are diagonal.

- **Example 3.**

When Σ_1 and Σ_2 are diagonal, we have a problem that can be decoupled into two problems similar to Case One. Consider finding the proper weights for two vector observations with the following *error structure*. Let

$$\Sigma_1 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}, \text{ and}$$

$$\Sigma_2 = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix} .$$

Both error structures are diagonal so the matrix problem will be broken into two separate scalar estimation problems (as in Case One). In the first problem, $\sigma_1 = 5$ and $\sigma_2 = 3$, yielding weights of 3/8 and 5/8. In the second, $\sigma_1 = 4$ and $\sigma_2 = 6$ giving weights of .6 and .4 to the observed value of the second variable. In matrix notation, the above process for finding the value to associate with the second vector observation can be written as follows:

$$K_2 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 10 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} .125 & 0 \\ 0 & .1 \end{bmatrix}$$

$$= \begin{bmatrix} .625 & 0 \\ 0 & .4 \end{bmatrix} .$$

This example demonstrates that new concepts are not needed when the covariance structure is diagonal.

The following observations simplify the calculations for the nondiagonal case. There is a theorem that states for any two covariance structures there is a basis in which they both are diagonal (Dempster 1969; Fukunaga 1972). Assume the appropriate change of basis is made. Then, using the method of Example Three, the appropriate weights can be found. A change of basis back to the original coordinate system will result in a matrix of full rank. We proceed to find a minimum variance estimator as follows:

$$E\left[(K_1 Z_1 + K_2 Z_2 - X)^2\right] = E\left[(K_1 V_1 + K_2 V_2)^2\right].$$

Since the different observations are uncorrelated, this can be written as

$$= E\left[(K_1 V_1)^2\right] + E\left[(K_2 V_2)^2\right]$$
$$= K_1 \Sigma_1 K'_1 + K_2 \Sigma_2 K'_2 .$$

In the above, K_2 can be expressed in terms of K_1

$$= K_1 \Sigma_1 K'_1 + (I - K_1) \Sigma_2 (I - K_1)' .$$

The error in this case is represented by a covariance matrix. Minimizing the trace of this covariance matrix will minimize the total estimation error. The rules for manipulating the trace of a matrix are discussed by Athans (1965). From his summary, the rule

$$\frac{\partial \operatorname{trace}(AXBX')}{\partial X} = A'XB' + AXB$$

is used letting A be the identity matrix, B be the covariance matrix, and X be K_1. The minimum occurs where

$$K_1 (\Sigma_1 + \Sigma_2) = \Sigma_2 ;$$

thus,

$$K_1 + \Sigma_2 (\Sigma_1 + \Sigma_2)^{-1},$$

and

$$K_2 = \Sigma_1 (\Sigma_1 + \Sigma_2)^{-1}.$$

The general vector solution can be written as a recursive estimator

$$\begin{aligned}
\hat{X}_{i+1} &= K_i \hat{X}_1 + K_{i+1} Z_{i+1} \\
&= \Sigma_{Z_{i+1}} (\Sigma_{\hat{X}_{i+1}} + \Sigma_{Z_{i+1}})^{-1} \hat{X}_1 + \Sigma_{\hat{X}_i} (\Sigma_{\hat{X}_i} + \Sigma_{Z_{i+1}})^{-1} Z_{i+1} \\
&= \hat{X}_i + \Sigma_{\hat{X}_i} (\Sigma_{\hat{X}_i} + \Sigma_{Z_{i+1}})^{-1} (Z_{i+1} - \hat{X}_i).
\end{aligned}$$

This should be interpreted as follows: the new estimate is the old estimate plus the product of the value of the new information, and the distance between the old estimate and the new observation. The gradient is the change to the old estimate. In many instances, the true error structure is not known. In these situations, it is sometimes possible to derive performance models from domain specific knowledge (usually based on analytic models of the system). These models can then be used to predict the error structure associated with a specific observation. Under different sets of assumptions, there are analytic and heuristic methods for finding a gradient.

- **Example 4.**

This example demonstrates the method for recursive estimation of a location in the x-y plane.

$$\text{Let } \hat{X}_1 = \begin{pmatrix} 100 \\ 100 \end{pmatrix}, \text{ and } \Sigma_{\hat{X}_i} = \begin{bmatrix} 10 & 5 \\ 5 & 30 \end{bmatrix}.$$

The correlation between successive estimates is assumed to be negligible.

Let the new observation be $Z_{i+1} = \begin{pmatrix} 95 \\ 110 \end{pmatrix}$, with covariance matrix $\Sigma_{Z_{i+1}} = \begin{bmatrix} 15 & 10 \\ 10 & 15 \end{bmatrix}$.

The new estimate $\hat{X}_{i+1}$ is:

$$\hat{X}_{i+1} = \begin{pmatrix} 100 \\ 100 \end{pmatrix} + \begin{bmatrix} 10 & 5 \\ 5 & 30 \end{bmatrix} \left[\begin{bmatrix} 10 & 5 \\ 5 & 30 \end{bmatrix} + \begin{bmatrix} 15 & 10 \\ 10 & 15 \end{bmatrix} \right]^{-1} \left[\begin{pmatrix} 95 \\ 110 \end{pmatrix} - \begin{pmatrix} 100 \\ 100 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 100 \\ 100 \end{pmatrix} + \begin{pmatrix} 10 & 5 \\ 5 & 30 \end{pmatrix} \begin{pmatrix} 45 & -15 \\ -15 & 25 \end{pmatrix} \frac{1}{900} \begin{pmatrix} -5 \\ 10 \end{pmatrix}$$

$$= \begin{pmatrix} 97.6389 \\ 108.75 \end{pmatrix}.$$

11. CONCLUSION

Consider again the estimation problem of a system that protects a tank from an incoming projectile. One Markov representation of the projectile is its trajectory in the X, Y, and Z dimensions. If there is more than one sensor estimating the location of the projectile at a given time step, then these should be combined into a single location estimate. If it is acceptable to assume a projectile travels at a constant velocity over the last portion of its path, the following three equations can be used to independently estimate the projectile trajectory in each dimension.

$$X = \alpha_1 + \alpha_2 t$$

$$Y = \beta_1 + \beta_2 t$$

$$Z = \gamma_1 + \gamma_2 t$$

Note that independence means parallel computations are possible. Each of the above can be solved using recursive least-squares processing. Improvements in the estimation are made by using recursive weighted least-squares estimation where the weights are based on estimates of the error in each dimension. To utilize all the information of the covariance structure of each estimate requires "en bloc" type updating using six by six matrices at each step of a recursive weighted least-squares process. After evaluating the available techniques, a final decision can be based on concerns of computation speed and acceptable accuracy of the estimator.

The wide array of uses for least-squares estimation testifies to its effectiveness. The key to structuring a problem for a least-squares solution is finding a Markov representation of the problem. This representation defines a recursive approach to estimation. When multiple estimates are available at each time step, the processing time can be decreased by using data fusion networks to reduce the information to a single estimate. Hierarchical networks using both parallel and serial combination of data can be devised. The cases presented herein can be used to preprocess the data for any recursive least-squares method.

12. REFERENCES

Athans, M. "Gradient Matrices and Matrix Calculations." Lincoln Laboratory, 1965 (AD-624426).

Dempster, A. P. Elements of Continuous Multivariate Analysis. Addison Wesley, 1969.

Fukunaga, K. Introduction to Statistical Pattern Recognition. Academic Press, 1972.

INTENTIONALLY LEFT BLANK.

No. of
Copies Organization

2 Administrator
Defense Technical Info Center
ATTN: DTIC-DDA
Cameron Station
Alexandria, VA 22304-6145

1 Commander
U.S. Army Materiel Command
ATTN: AMCAM
5001 Eisenhower Avenue
Alexandria, VA 22333-0001

1 Commander
U.S. Army Laboratory Command
ATTN: AMSLC-DL
2800 Powder Mill Road
Adelphi, MD 20783-1145

2 Commander
U.S. Army Armament Research,
Development, and Engineering Center
ATTN: SMCAR-IMI-I
Picatinny Arsenal, NJ 07806-5000

2 Commander
U.S. Army Armament Research,
Development, and Engineering Center
ATTN: SMCAR-TDC
Picatinny Arsenal, NJ 07806-5000

1 Director
Benet Weapons Laboratory
U.S. Army Armament Research,
Development, and Engineering Center
ATTN: SMCAR-CCB-TL
Watervliet, NY 12189-4050

(Unclass. only)1 Commander
U.S. Army Armament, Munitions
and Chemical Command
ATTN: AMSMC-IMF-L
Rock Island, IL 61299-5000

1 Director
U.S. Army Aviation Research
and Technology Activity
ATTN: SAVRT-R (Library)
M/S 219-3
Ames Research Center
Moffett Field, CA 94035-1000

No. of
Copies Organization

1 Commander
U.S. Army Missile Command
ATTN: AMSMI-RD-CS-R (DOC)
Redstone Arsenal, AL 35898-5010

1 Commander
U.S. Army Tank-Automotive Command
ATTN: ASQNC-TAC-DIT (Technical
Information Center)
Warren, MI 48397-5000

1 Director
U.S. Army TRADOC Analysis Command
ATTN: ATRC-WSR
White Sands Missile Range, NM 88002-5502

1 Commandant
U.S. Army Field Artillery School
ATTN: ATSF-CSI
Ft. Sill, OK 73503-5000

(Class. only)1 Commandant
U.S. Army Infantry School
ATTN: ATSH-CD (Security Mgr.)
Fort Benning, GA 31905-5660

(Unclass. only)1 Commandant
U.S. Army Infantry School
ATTN: ATSH-CD-CSO-OR
Fort Benning, GA 31905-5660

1 Air Force Armament Laboratory
ATTN: WL/MNOI
Eglin AFB, FL 32542-5000

Aberdeen Proving Ground

2 Dir, USAMSAA
ATTN: AMXSY-D
AMXSY-MP, H. Cohen

1 Cdr, USATECOM
ATTN: AMSTE-TC

3 Cdr, CRDEC, AMCCOM
ATTN: SMCCR-RSP-A
SMCCR-MU
SMCCR-MSI

1 Dir, VLAMO
ATTN: AMSLC-VL-D

10 Dir, BRL
ATTN: SLCBR-DD-T

No. of Copies	Organization
1	Director U.S. Army Aviation Research and Technology Activity ATTN: SAVRT-R (Library) M/S 219-3 Ames Research Center Moffett Field, CA 94035-1000
1	Director NSA ATTN: ESI, Charles Alexander 9800 Savage Road Ft. Meade, MD 20775
1	University of Maryland Baltimore County Department of Mathematics and Statistics ATTN: Bimal K. Sinha 5401 Wilkens Avenue Catonsville, MD 21228
1	M.I.T. Center for Intelligent Control Systems ATTN: J. N. Tsitsiklis Bldg 35-214 Cambridge, MA 02139

No. of Copies	Organization
	Aberdeen Proving Ground
9	Dir, USAMSAA ATTN: AMXSY-AA AMXSY AD AMXSY-CA, Paul Kunselman AMXSY-CA, William Clay AMXSY-CA, Pat O'Neill AMXSY-CC, Richard Sandmeyer AMXSY-GA AMXSY-GS AMXSY-RM, John Woodworth

No. of Copies	Copies
1	McMaster University Department of Electrical and Computer Engineering ATTN: Z. Q. Luo Room 225/GRL Hamilton, Ontario, L8S 4L7 Canada

INTENTIONALLY LEFT BLANK.

USER EVALUATION SHEET/CHANGE OF ADDRESS

This laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers below will aid us in our efforts.

1. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) ____________________

2. How, specifically, is the report being used? (Information source, design data, procedure, source of ideas, etc.) ____________________

3. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. ____________________

4. General Comments. What do you think should b- changed to improve future reports? (Indicate changes to organization, technical content, fo.mat, etc.) ____________________

BRL Report Number BRL-TR-3303 Division Symbol ____________

Check here if desire to be removed from distribution list. ____

Check here for address change. ____

Current address: Organization ____________________
Address ____________________

DEPARTMENT OF THE ARMY
Director
U.S. Army Ballistic Research Laboratory
ATTN: SLCBR-DD-T
Aberdeen Proving Ground, MD 21005-5066

OFFICIAL BUSINESS

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT No 0001, APG, MD

Postage will be paid by addressee

Director
U.S. Army Ballistic Research Laboratory
ATTN: SLCBR-DD-T
Aberdeen Proving Ground, MD 21005-5066